

What is claimed is:

1. An object activity modeling method comprising the steps of:

- (a) obtaining an optical flow vector from a video sequence;
- (b) obtaining the probability distribution of the feature vector for  
5 a plurality of video frames, using the optical flow vector;
- (c) modeling states, using the probability distribution of the feature vector; and
- (d) expressing the activity of the object in the video sequence based on state transition.

2. The object activity modeling method of claim 1, wherein the step (a) is based on affine motion estimation.

3. The object activity modeling method of claim 2, wherein the step (a) further comprises the sub-steps of:

- (a-1) grouping input video frames into a plurality of video frame groups and dividing each video frame group as an individual state;
- 5 (a-2) obtaining an affine motion parameter for each video in the video frame group of each individual state; and
- (a-3) obtaining an optical flow vector from the affine motion parameters.

4. The object activity modeling method of claim 3, wherein the step (a-2) comprises a step for determining parameters, which minimizes summed square difference  $\sum (I_t(x) - I_{t-1}(x - V(x)))^2$  over a given video based on the intensity of the pixel on the object, which is expressed as  $I_t(x) = I_{t-1}(x - V(x))$  when  $I$  denotes intensity,  $t$  denotes time,  $x$  denotes a pixel location  $(x, y)$ , and  $v$  denotes the motion vector, as motion parameters.

5. The object activity modeling method of claim 1, wherein the step (b) comprises a step for calculating probability distribution  $P(Z|\Omega)$  by the following equation:

$$P(Z|\Omega) = \frac{\exp(-\frac{1}{2}(z - m)^T Q^{-1}(Z - m))}{(2\pi)^N |Q|^{1/2}}$$

wherein  $P = (p_1, p_2, \dots, p_d)$  denotes a motion vector calculated at each pixel location  $(x, y)$ ,  $L$  denotes the number of pixels in a video frame or a region of interest,  $d$  denotes the number of dimensions, feature vector  $Z$ , which is a  $d \times L$  dimension vector, is  $Z = (p_1^1, p_1^2, \dots, p_1^L, p_2^1, p_2^2, \dots, p_2^L, p_d^1, p_d^2, \dots, p_d^L)^T$ ,  $m$  is the mean vector of feature vector  $Z$ , and  $Q$  is the covariance matrix of feature vector  $Z$ , and it is assumed that feature vector  $Z$  is provided from observation class  $\Omega$ .

6. The object activity modeling method of claim 1, wherein the step (b) further comprises the steps of:

decomposing covariance matrix Q as the following equation:

$$Q = \Phi \Lambda \Phi^T$$

5       Wherein  $\hat{Z}$  is equal to Z-m, the columns of  $\Phi$  are orthonormal eigenvectors of covariance matrix Q, and A corresponds to the diagonal eigenvalue; and

calculating probability distribution  $P(Z|\Omega)$  by the following equation:

$$P(Z|\Omega) = \left[ \frac{\exp(-\frac{1}{2} \sum_i^M y_i^2 / \alpha_i)}{(2\pi)^M |\Lambda|^{1/2}} \right] \left[ \frac{\exp(-\frac{1}{2} \sum_{M+1}^N y_i^2 / 2\rho)}{(2\pi\rho)^{(N-M)/2}} \right]$$

10       wherein M is the number of principal components,  $y_i$  is the i-th component of Y,  $\alpha_i$  is the i-th eigenvalue of Q, and  $\rho$  is the optimal

value, which is obtained by  $\rho = \frac{1}{N-M} \sum_{M+1}^N \alpha_i$ , and it is assumed that

feature vector Z is provided from observation class  $\Omega$ .

7. The object activity modeling method of claim 1, wherein in the step (c), the object activity in the video sequence is expressed using a Hidden Markov Model (HMM), based on state transition.

8. The object activity modeling method of claim 7, wherein the Hidden Markov Model (HMM) is expressed as  $\lambda = \{\Xi, A, B, \Pi\}$  when  $N$  is the number of possible states,  $\Xi$  satisfies  $\Xi = \{q_1, q_2, \dots, q_N\}$ ,  $A$  is  $\{a_{ij}\}$ , the transition between hidden states  $i$  and  $j$ ,  $B$  is  $\{b_j(\cdot)\}$ , the observation symbol probability corresponding to state  $j$ , and  $\Pi$  is the initial state distribution, and the state  $\Xi = \{q_1, q_2, \dots, q_N\}$  and the initial state distribution  $\Pi$  are determined in advance based on video data.

9. An object activity recognition method comprising the steps of:

(a) obtaining feature vectors by motion estimation for video frames;

(b) determining a state, to which each frame belongs, using the obtained feature vectors; and

(c) determining an activity model, which maximizes the probability between activity models and a video frame provided from a given activity model dictionary using a transition matrix for the determined state, as the recognized activity.

10. The object activity recognition method of claim 9, wherein the step (d) comprises a step of finding an activity model, which maximizes probability  $P(O|\lambda)$  from the given activity model dictionary  $\{\lambda_1, \lambda_2, \dots, \lambda_E\}$ , when  $T$  is a positive integer indicating the number of

- 5 frames forming the video sequence,  $Z_1, Z_2, \dots, Z_T$  are feature vectors of first frame, second frame, ..., T-th frame, respectively, and if video frame  $O=\{Z_1, Z_2, \dots, Z_T\}$  is given and  $E$  is the number of state models.

11. The object activity recognition method of claim 10, wherein the transition matrix is obtained by using an expectation-maximization (EM) algorithm based on the observation symbol probability  $\{b_j(\cdot)\}$  corresponding to scene  $j$  in the training process.